

7



ENTERED

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/029,840

DATE: 04/12/2002

TIME: 14:27:57

Input Set : A:\EP.txt

Output Set: N:\CRF3\04122002\J029840.raw

3 <110> APPLICANT: Meng, Xiang-Jin
4 Haqshenas, Gholamreza
5 Huang, Fang-Fang
7 <120> TITLE OF INVENTION: Avian Hepatitis E Virus, Vaccines and Methods of Protecting
Against Avian
8 Hepatitis-Spenomegaly Syndrome and Mammalian Hepatitis E
10 <130> FILE REFERENCE: AM100389
12 <140> CURRENT APPLICATION NUMBER: US 10/029,840
13 <141> CURRENT FILING DATE: 2001-12-31
15 <160> NUMBER OF SEQ ID NOS: 11
17 <170> SOFTWARE: PatentIn version 3.1
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 3946
21 <212> TYPE: DNA
22 <213> ORGANISM: Hepatitis E virus
24 <400> SEQUENCE: 1
25 accagcattg gatttcgatg gacgctgttt aacgagcgcc gttgatcttg ggttgccagcc 60
27 taccagctgg cgcaccgtat cccaccgttg cccttgggac gtttgatat ttttgcgtac 120
29 tgattatccg actatcacca caaccagtag ggtgctgcgg tctgttggtg ttaccgggtga 180
31 aaccattggt cagaagatag tgtttaccca ggtggccaag cagtcgaacc cggggtccat 240
33 aacgggtccat gaggcgcagg gcagtacttt tgatcagact actataatcg ccacgttaga 300
35 tgctcgtggc cttatagctt catctcgcgc gcatgccata gttgcgctaa cccgccaccg 360
37 ggagcgtgtg agtgtgattg atgttggtgg ggtgctggtc gagattggag ttactgatgc 420
39 catgtttaac aatatcgaaa tgcagcttgt gcgacctgat gctgcagccc ctgccggggt 480
41 gctacgagcc ccagacgaca ccgtggatgg cttgttggac atacccccgg cccacactga 540
43 tgtagcggcg gtgttaacag ctgaggcgat tgggcatgcg ccccttgaat tggccgccat 600
45 aaatccaccc gggcctgtat tggagcaggg cctattatac atgccggcca ggcttgatgg 660
47 gcgtgatgag gttgttaagc tccagctgtc ggatactgta cactgcgcgc tggctgcacc 720
49 cactagccgt cttgcgggtga ttaacacatt ggttgggagg tacggtaaag ccactaagct 780
51 gctgaggtt gaatatgact taatggacac tattgcgcag ttctggcatc atatcggacc 840
53 aatcaacccc tcaacactgg agtatgcaga gatgtgcgag gccatgotta gtaagggccca 900
55 ggatgggtcc ttgattgtac atctggattt acaggatgct gattgttctc gcataacatt 960
57 cttccagaag gactgcgcta aatttacgct ggatgacct gttgcacaag gtaaagtggg 1020
59 acaggggata tctgcgtggc cgaaaacttt gtgtgcactt ttccggccct ggttccgggc 1080
61 tatagagaag caccttgttg ctgggttacc cccaggttat tactatgggg acctgtacac 1140
63 ggaagccgat ctgcatcgtt ctgtgctttg cgcgcctgct ggtcaccttg tttttgagaa 1200
65 tgatttctca gagtgttgact caacgcagaa taattgttcc cttgatctcg aatgtgaatt 1260
67 gatgcgcagg tttgggatgc ccgattggat ggtagccttg taccatcttg ttcgatcata 1320
69 ctggctcttg gttgccccga aagaagccct tctgtgctgt tggaaaaaac actctggtga 1380
71 gccgggcacc cttttgtgga atacagtttg gaacatgact gtgttgcatc atgtttatga 1440
73 gtttgatcga ccaagtgtgt tgtgtttcaa aggtgatgat agtgtcgttg tctgtgaatc 1500
75 ggtgcgcgcc cgtccagagg gcgttagtct cgtggcagac tgcgggctaa aaatgaagga 1560
77 caagaccggc ccgtgtggcg ctttttccaa cctgctgata ttcccgggag ctggtgttgt 1620
79 ctgcgacctg ttacggcagt ggggcccgtt gactgacaag aactgggggc ccgacattca 1680

RAW SEQUENCE LISTING

DATE: 04/12/2002

PATENT APPLICATION: US/10/029,840

TIME: 14:27:57

Input Set : A:\EP.txt

Output Set: N:\CRF3\04122002\J029840.raw

```

81 gcggatgcag gaccttgagc aagcgtgtaa ggattttgtt gcacgtgttg taactcaggg 1740
83 taaagagatg ttgaccatcc agcttgtggc gggttattat ggtgtggaag ttggtatggt 1800
85 tgagggtggt tggggggctt tgaaggcctg cgccgcagcc cgcgagaccc tagtgaccaa 1860
87 caggttgccg gtactaaact tatctaagga ggactgaaca aataacaatc attatgcagt 1920
89 ctgcgcgtcc atgtgcctta gctgccagtt ctgggtgttt gagtgccagg aaagtggggg 1980
91 gggatgtcgc tgtgtagatt gttgctcatg cttgcaatgt gctgcggggg gtcaaggggc 2040
93 tcccaaacgc tccagccggg aggcaggcgt ggccagcgcc gccgtgacaa ttcagcccag 2100
95 tggagcactc aacaacgccc cgaggagccc gtcggccccc cccctctcac agacgttgtc 2160
97 accgcggcag gtactcgcac ggtaccagat gtagatcaag ccggtgccgt gctggtgcgc 2220
99 cagtataatc tagtgaccag cccgttaggc ctggccaccc ttggtagcac caatgccttg 2280
101 ctttatgccg caccggtgtc accgttaatg ccgcttcagg acggcacgac gtctaataatc 2340
103 atgagcacgg agtctagcaa ctatgctcaa taccgtgtac agggcctaac tgtccgctgg 2400
105 cgcccagttg tgccaaatgc ggtgggcggc ttctctataa gcatggccta ttggccccag 2460
107 acaacatcca cccctacaag cattgacatg aattccatca cgtccactga cgtccgtgtg 2520
109 gtgcttcagc cgggctctgc tggtttgcgt actataccac atgagcgttt ggcgtataag 2580
111 aacaatggtt ggcggtccgt cgaaacggta tccgtcccac aggaggatgc cacgtccggc 2640
113 atgctcatgg tttgtgtcca cgggaccccc tggaatagtt ataccaatag tgtttacacc 2700
115 gggccgcttg gtatggttga ttttgccata aagttacagc taaggaaactt gtcgcccggt 2760
117 aatacaaatg ccagggtcac ccgtgtgaag gtgacggccc cacataccat caaggctgac 2820
119 ccactctggt ctaccataac aacagcagct gcggccagggt ttatggcgga tgtgcgttgg 2880
121 ggcttgggca ctgctgagga tggcgaaatt ggtcacggca tccttgggtg tctgtttaac 2940
123 ctggcgga caagttttagg tggcttggcc tcgacactgc tgcgggcggc gagggttcag 3000
125 tacatgtacg gccggcctgt ggggaacgcg aacggcgagc ctgaggtgaa actgtatatg 3060
127 tcggttgagg atgcggttaa cgataaacct attatggtcc cccatgacat cgacctcggg 3120
129 accagcactg tcacctgcca ggactatggg aatcagcatg tggatgaccg cccatccccg 3180
131 gccccggccc ctaagcgagc tttgggcacc ctaaggctcag gggatgtgtt gcgtattact 3240
133 ggctccatgc agtatgtgac taacgcccag ttgttaccgc agagtgtgtc acaggggtac 3300
135 tttggggccg gcagcaccat gatggtgcat aatttgatca ctggtgtgcg cgcccccgcc 3360
137 agttcagtcg actggacgaa ggcaacagtg gatgggtcc aggtgaagac tgtcgatgct 3420
139 agttctggga gtaatagggt tgcagcggtt cctgcatttg gaaagccagc tgtgtggggg 3480
141 cccagggcg ctgggtatit ctaccagtat aacagcacc accaggagtg gatttatit 3540
143 cttcagaatg gtagctccgt ggtttggtat gcataacta atatgttggg ccagaagtca 3600
145 gatacatcca ttctttttga ggtccggcca atccaagcta gtgatcagcc ttggtttttg 3660
147 gcacaccaca ctggcgggga tgactgtacc acctgtctgc ctctgggggt aagaacatgt 3720
149 tgccgccagg cgccagaaga ccagtcacct gagacgcgcc ggctcctaga ccggttagt 3780
151 aggacattcc cctcaccacc ctaatgtcgt ggttttgggg ttttaggttg attttctgta 3840
153 tctgggcgta attgccccta tgtttaaitt attgtgattt ttataactgt tcatttgatt 3900
155 atttatgaaa tcctcccatc tcgggcatag taaaaaaaa aaaaaa 3946

```

158 <210> SEQ ID NO: 2

159 <211> LENGTH: 146

160 <212> TYPE: PRT

161 <213> ORGANISM: Hepatitis E virus

163 <400> SEQUENCE: 2

165 Pro Ala Leu Asp Phe Asp Gly Arg Cys Leu Thr Ser Ala Val Asp Leu

166 1 5 10 15

169 Gly Leu Gln Pro Thr Ser Trp Arg Thr Val Ser His Arg Cys Pro Trp

170 20 25 30

173 Asp Val Cys Ile Phe Leu Arg Thr Asp Tyr Pro Thr Ile Thr Thr Thr

174 35 40 45

RAW SEQUENCE LISTING

DATE: 04/12/2002

PATENT APPLICATION: US/10/029,840

TIME: 14:27:57

Input Set : A:\EP.txt

Output Set: N:\CRF3\04122002\J029840.raw

```

177 Ser Arg Val Leu Arg Ser Val Val Phe Thr Gly Glu Thr Ile Gly Gln
178      50                      55                      60
181 Lys Ile Val Phe Thr Gln Val Ala Lys Gln Ser Asn Pro Gly Ser Ile
182 65                      70                      75                      80
185 Thr Val His Glu Ala Gln Gly Ser Thr Phe Asp Gln Thr Thr Ile Ile
186                      85                      90                      95
189 Ala Thr Leu Asp Ala Arg Gly Leu Ile Ala Ser Ser Arg Ala His Ala
190                      100                      105                      110
193 Ile Val Ala Leu Thr Arg His Arg Glu Arg Cys Ser Val Ile Asp Val
194                      115                      120                      125
197 Gly Gly Val Leu Val Glu Ile Gly Val Thr Asp Ala Met Phe Asn Asn
198      130                      135                      140
201 Ile Glu
202 145
205 <210> SEQ ID NO: 3
206 <211> LENGTH: 439
207 <212> TYPE: DNA
208 <213> ORGANISM: Hepatitis E virus
210 <400> SEQUENCE: 3
211 accagcattg gatttcgatg gacgctgttt aacgagcgcc gttgatcttg ggttgcagcc      60
213 taccagctgg cgcaccgtat cccaccgttg cccttgggac gtttgtatat ttttgcgtac      120
215 tgattatccg actatcacca caaccagtag ggtgctgcgg tctgttggtg ttaccggtga      180
217 aaccattggt cagaagatag tgtttaccca ggtggccaag cagtcgaacc ccgggtccat      240
219 aacgggtccat gaggcgcagg gcagtaacttt tgatcagact actataatcg ccacgttaga      300
221 tgctcgtggc cttatagctt catctcgcgc gcatgccata gttgcgctaa cccgccaccg      360
223 ggagcgctgt agtgtgattg atgttggttg ggtgctggtc gagattggag ttactgatgc      420
225 catgtttaac aatatcgaa                                     439
228 <210> SEQ ID NO: 4
229 <211> LENGTH: 483
230 <212> TYPE: PRT
231 <213> ORGANISM: Hepatitis E virus
233 <400> SEQUENCE: 4
235 Leu Val Arg Pro Asp Ala Ala Ala Pro Ala Gly Val Leu Arg Ala Pro
236 1      5      10      15
239 Asp Asp Thr Val Asp Gly Leu Leu Asp Ile Pro Pro Ala His Thr Asp
240      20      25      30
243 Val Ala Ala Val Leu Thr Ala Glu Ala Ile Gly His Ala Pro Leu Glu
244      35      40      45
247 Leu Ala Ala Ile Asn Pro Pro Gly Pro Val Leu Glu Gln Gly Leu Leu
248      50      55      60
251 Tyr Met Pro Ala Arg Leu Asp Gly Arg Asp Glu Val Val Lys Leu Gln
252 65      70      75      80
255 Leu Ser Asp Thr Val His Cys Arg Leu Ala Ala Pro Thr Ser Arg Leu
256      85      90      95
259 Ala Val Ile Asn Thr Leu Val Gly Arg Tyr Gly Lys Ala Thr Lys Leu
260      100     105     110
263 Pro Glu Val Glu Tyr Asp Leu Met Asp Thr Ile Ala Gln Phe Trp His
264      115     120     125
267 His Ile Gly Pro Ile Asn Pro Ser Thr Leu Glu Tyr Ala Glu Met Cys

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/029,840

DATE: 04/12/2002

TIME: 14:27:57

Input Set : A:\EP.txt

Output Set: N:\CRF3\04122002\J029840.raw

```

268      130      135      140
271 Glu Ala Met Leu Ser Lys Gly Gln Asp Gly Ser Leu Ile Val His Leu
272 145      150      155      160
275 Asp Leu Gln Asp Ala Asp Cys Ser Arg Ile Thr Phe Phe Gln Lys Asp
276      165      170      175
279 Cys Ala Lys Phe Thr Leu Asp Asp Pro Val Ala His Gly Lys Val Gly
280      180      185      190
283 Gln Gly Ile Ser Ala Trp Pro Lys Thr Leu Cys Ala Leu Phe Gly Pro
284      195      200      205
287 Trp Phe Arg Ala Ile Glu Lys His Leu Val Ala Gly Leu Pro Pro Gly
288      210      215      220
291 Tyr Tyr Tyr Gly Asp Leu Tyr Thr Glu Ala Asp Leu His Arg Ser Val
292 225      230      235      240
295 Leu Cys Ala Pro Ala Gly His Leu Val Phe Glu Asn Asp Phe Ser Glu
296      245      250      255
299 Phe Asp Ser Thr Gln Asn Asn Val Ser Leu Asp Leu Glu Cys Glu Leu
300      260      265      270
303 Met Arg Arg Phe Gly Met Pro Asp Trp Met Val Ala Leu Tyr His Leu
304      275      280      285
307 Val Arg Ser Tyr Trp Leu Leu Val Ala Pro Lys Glu Ala Leu Arg Gly
308      290      295      300
311 Cys Trp Lys Lys His Ser Gly Glu Pro Gly Thr Leu Leu Trp Asn Thr
312 305      310      315      320
315 Val Trp Asn Met Thr Val Leu His His Val Tyr Glu Phe Asp Arg Pro
316      325      330      335
319 Ser Val Leu Cys Phe Lys Gly Asp Asp Ser Val Val Val Cys Glu Ser
320      340      345      350
323 Val Arg Ala Arg Pro Glu Gly Val Ser Leu Val Ala Asp Cys Gly Leu
324      355      360      365
327 Lys Met Lys Asp Lys Thr Gly Pro Cys Gly Ala Phe Ser Asn Leu Leu
328      370      375      380
331 Ile Phe Pro Gly Ala Gly Val Val Cys Asp Leu Leu Arg Gln Trp Gly
332 385      390      395      400
335 Arg Leu Thr Asp Lys Asn Trp Gly Pro Asp Ile Gln Arg Met Gln Asp
336      405      410      415
339 Leu Glu Gln Ala Cys Lys Asp Phe Val Ala Arg Val Val Thr Gln Gly
340      420      425      430
343 Lys Glu Met Leu Thr Ile Gln Leu Val Ala Gly Tyr Tyr Gly Val Glu
344      435      440      445
347 Val Gly Met Val Glu Val Val Trp Gly Ala Leu Lys Ala Cys Ala Ala
348      450      455      460
351 Ala Arg Glu Thr Leu Val Thr Asn Arg Leu Pro Val Leu Asn Leu Ser
352 465      470      475      480
355 Lys Glu Asp
359 <210> SEQ ID NO: 5
360 <211> LENGTH: 1450
361 <212> TYPE: DNA
362 <213> ORGANISM: Hepatitis E virus
364 <400> SEQUENCE: 5

```

RAW SEQUENCE LISTING

DATE: 04/12/2002

PATENT APPLICATION: US/10/029,840

TIME: 14:27:57

Input Set : A:\EP.txt

Output Set: N:\CRF3\04122002\J029840.raw

```

365 gcttgtgcga cctgatgctg cagcccctgc cggggtgcta cgagccccag acgacacogt      60
367 ggatggcttg ttggacatac ccccggccca cactgatgta gcggcggtgt taacagctga      120
369 ggcgattggg catgcgcccc ttgaattggc cgccataaat ccacccgggc ctgtattgga      180
371 gcagggccta ttatacatgc cggccaggct tgatgggcgt gatgagggtt ttaagctcca      240
373 gctgtcggat actgtacact gccgcctggc tgcacccact agccgtcttg cggtgattaa      300
375 cacattggtt ggcgcgtacg gtaaagccac taagctgcct gaggttgaat atgacttaat      360
377 ggacactatt gcgcagttct ggcatcatat cggaccaatc aacccctcaa cactggagta      420
379 tgcagagatg tgcgaggcca tgcttagtaa gggccaggat gggtccttga ttgtacatct      480
381 ggattttacag gatgctgatt gttctcgcat aacattcttc cagaaggact gcgctaaatt      540
383 tacgctggat gacctgttg cacacggtaa agtgggacag gggatatctg cgtggccgaa      600
385 aacttttgtt gcacttttcg gccctggtt ccgggctata gagaagcacc ttgtggctgg      660
387 gttacccccca ggttattact atggggacct gtacacggaa gccgatctgc atcgttctgt      720
389 gctttgcgcg cctgctggtc acctgtttt tgagaatgat ttctcagagt ttgactcaac      780
391 gcagaataat gtgtcccttg atctcgaatg tgaattgatg cgcaggtttg ggatgcccgga      840
393 ttggatggta gccttgtaac atcttgttcg atcatactgg ctcttggttg ccccgaaaga      900
395 agcccttcgt ggctgttgga aaaaacactc tggtgagccg ggcacccttt tgtggaatac      960
397 agtttggaac atgactgtgt tgcacatgtt ttatgagttt gatcgaccaa gtgtgttgtg      1020
399 tttcaaaggt gatgatagtg tcgttgtctg tgaatcgggt gcgcgccgtc cagagggcgt      1080
401 tagtctcgtg gcagactgcg ggctaaaaat gaaggacaag accggcccggt gtggcgccctt      1140
403 ttccaacctg ctgatcttcc cgggagctgg tgttgtctgc gacctgttac ggcagtgggg      1200
405 ccgcttgact gacaagaact gggggcccgga cattcagcgg atgcaggacc ttgagcaagc      1260
407 gtgtaaggat tttgttgcaac gtgttgtaac tcagggtaaa gagatgttga ccatccagct      1320
409 tgtggcggtt tattatggtg tggaagtggg tatggttgag gtggtttggg gggctttgaa      1380
411 ggcctgcgcg gcagcccgcg agaccctagt gaccaacagg ttgccggtac taaacttate      1440
413 taaggaggac                                     1450
416 <210> SEQ ID NO: 6
417 <211> LENGTH: 606
418 <212> TYPE: PRT
419 <213> ORGANISM: Hepatitis E virus
421 <400> SEQUENCE: 6
423 Met Ser Leu Cys Arg Leu Leu Leu Met Leu Ala Met Cys Cys Gly Val
424 1          5          10          15
427 Ser Arg Gly Ser Gln Thr Leu Pro Ala Gly Gly Arg Arg Gly Gln Arg
428          20          25          30
431 Arg Arg Asp Asn Ser Ala Gln Trp Ser Thr Gln Gln Arg Pro Glu Gly
432          35          40          45
435 Ala Val Gly Pro Ala Pro Leu Thr Asp Val Val Thr Ala Ala Gly Thr
436          50          55          60
439 Arg Thr Val Pro Asp Val Asp Gln Ala Gly Ala Val Leu Val Arg Gln
440 65          70          75          80
443 Tyr Asn Leu Val Thr Ser Pro Leu Gly Leu Ala Thr Leu Gly Ser Thr
444          85          90          95
447 Asn Ala Leu Leu Tyr Ala Ala Pro Val Ser Pro Leu Met Pro Leu Gln
448          100         105         110
451 Asp Gly Thr Thr Ser Asn Ile Met Ser Thr Glu Ser Ser Asn Tyr Ala
452          115         120         125
455 Gln Tyr Arg Val Gln Gly Leu Thr Val Arg Trp Arg Pro Val Val Pro
456          130         135         140
459 Asn Ala Val Gly Gly Phe Ser Ile Ser Met Ala Tyr Trp Pro Gln Thr

```

VERIFICATION SUMMARY

DATE: 04/12/2002

PATENT APPLICATION: US/10/029,840

TIME: 14:27:58

Input Set : A:\EP.txt

Output Set: N:\CRF3\04122002\J029840.raw